KMAN PUBLISHING INSTITUTE
Journal Website

Article history:
Received 05 November 2025
Revised 13 February 2026
Accepted 21 February 2026
Published online 01 May 2026

International Journal of Innovation Management and Organizational Behavior

Volume 6, Issue 3, pp 1-13

IJIMOB
International Journal of Innovation Management and Organizational Behavior

E-ISSN: 3041-8992

Predicting Employee Turnover Using Explainable Machine Learning Models: A Comparative Study of XGBoost, LightGBM, CatBoost, and Deep Neural Networks

Gabriela. Torres[1], Priya. Nair[2*], Mikko Lahtinen[3]

[1] Department of Business Administration and Innovation, Tecnológico de Monterrey, Monterrey, Mexico
[2] Department of Human Resource Management, Ted Rogers School of Management, Toronto Metropolitan University, Toronto, Canada
[3] Department of Industrial Engineering and Management, Aalto University, Espoo, Finland

*** Corresponding author email address**: priya.nair@torontomu.ca

Article Info

Article type:
Original Research

How to cite this article:
Torres, G., Nair, P., & Lahtinen, M. (2026). Predicting Employee Turnover Using Explainable Machine Learning Models: A Comparative Study of XGBoost, LightGBM, CatBoost, and Deep Neural Networks. *International Journal of Innovation Management and Organizational Behavior, 6*(3), 1-13.
https://doi.org/10.61838/kman.ijimob.5454

ABSTRACT

Objective: This study aimed to develop and compare explainable machine learning models for employee turnover prediction using XGBoost, LightGBM, CatBoost, and Deep Neural Networks and to identify the most influential predictors of turnover risk through explainable artificial intelligence techniques.

Methods and Materials: This quantitative predictive analytics study was conducted using a dataset of 4,862 employees from 37 organizations across multiple industries in Canada. Employee demographic, organizational, behavioral, performance-related, and psychological variables were collected from human resource information systems and validated survey instruments. The dataset was preprocessed through normalization, feature engineering, categorical encoding, and missing-value imputation procedures. Four machine learning algorithms, including XGBoost, LightGBM, CatBoost, and Deep Neural Networks, were developed and optimized using grid search and Bayesian hyperparameter tuning. The dataset was divided into training, validation, and testing subsets using stratified sampling. Model performance was evaluated using accuracy, precision, recall, F1-score, area under the receiver operating characteristic curve (AUC-ROC), area under the precision-recall curve (AUC-PR), Matthews correlation coefficient, and balanced accuracy. Explainability was achieved through SHapley Additive exPlanations (SHAP), Local Interpretable Model-Agnostic Explanations (LIME), and partial dependence analyses.

Findings: The comparative analysis revealed that all machine learning models achieved strong predictive performance; however, CatBoost outperformed the other algorithms, achieving an accuracy of 91.9%, an F1-score of 0.878, and an AUC-ROC value of 0.958. XGBoost demonstrated the second-highest performance with an accuracy of 91.3% and an AUC-ROC of 0.951, followed by LightGBM (accuracy = 90.6%, AUC-ROC = 0.944) and the Deep Neural Network (accuracy = 89.7%, AUC-ROC = 0.936). ROC curve analyses confirmed the superior discriminative ability of CatBoost. SHAP-based explainability analyses identified job satisfaction, employee engagement, organizational commitment, perceived organizational support, promotion frequency, salary growth, performance ratings, overtime hours, absenteeism, and training participation as the most influential predictors of employee turnover. The confusion matrix of the CatBoost model further demonstrated strong classification capability with high true positive and true negative rates and minimal classification errors.

Conclusion: The findings demonstrate that explainable machine learning models can accurately predict employee turnover and provide meaningful insights into the organizational, behavioral, and psychological factors associated with employee departure. Among the evaluated algorithms, CatBoost offered the optimal balance between predictive accuracy and interpretability. The integration of explainable artificial intelligence techniques enhanced transparency and practical applicability, enabling organizations to identify high-risk employees and implement proactive retention strategies. These results support the adoption of explainable machine learning as an effective decision-support tool for workforce analytics, talent management, and organizational sustainability.

Keywords: *Employee Turnover Prediction; Explainable Artificial Intelligence; Human Resource Analytics; Machine Learning*

1 Introduction

Employee turnover has emerged as one of the most significant challenges confronting contemporary organizations. In increasingly competitive labor markets, organizations invest substantial resources in recruiting, training, developing, and retaining talented employees. When employees voluntarily leave an organization, the consequences extend beyond the direct costs of recruitment and replacement. Turnover disrupts organizational routines, diminishes institutional knowledge, reduces team cohesion, negatively affects productivity, and may compromise organizational competitiveness. Consequently, understanding, predicting, and mitigating employee turnover has become a strategic priority for human resource management and organizational leadership. Traditional approaches to turnover analysis have generally relied on retrospective assessments, descriptive statistics, and theoretical models that explain why employees leave after turnover has already occurred. While these approaches have contributed substantially to the turnover literature, they often provide limited support for proactive intervention and real-time decision-making. The growing availability of organizational data and advances in computational technologies have created opportunities for organizations to transition from reactive to predictive workforce management. In this context, machine learning has emerged as a promising analytical paradigm capable of identifying complex patterns associated with employee turnover and generating actionable insights that support evidence-based retention strategies (Dhamija, 2023; Kiran et al., 2023; Patnaik, 2024).

The evolution of Human Resource Analytics (HR Analytics) has played a central role in transforming workforce management practices. HR Analytics integrates statistical techniques, data mining methods, and predictive algorithms to support strategic human resource decision-making. Rather than relying solely on intuition or managerial judgment, organizations increasingly utilize data-driven approaches to understand employee behavior, assess workforce risks, and optimize organizational outcomes. Employee turnover prediction represents one of the most important applications of HR Analytics because it directly influences organizational performance, workforce stability, and long-term sustainability. Recent studies have emphasized that predictive analytics can help organizations identify employees who are at risk of leaving before turnover actually occurs, enabling managers to implement targeted interventions and retention initiatives. Such interventions may include career development opportunities, compensation adjustments, leadership support programs, workload redistribution, and employee engagement initiatives. Consequently, predictive turnover modeling has become an essential component of modern talent management frameworks (Kundu et al., 2024; Ravesangar & Narayanan, 2024; Tripathi et al., 2024).

The rapid expansion of big data technologies has further accelerated interest in employee turnover prediction. Organizations now generate large volumes of workforce-related data through human resource information systems, performance management platforms, employee surveys, attendance records, communication systems, and digital collaboration tools. These data sources provide rich information regarding employee demographics, job characteristics, organizational experiences, behavioral patterns, and performance outcomes. The integration of big data with advanced analytical techniques has enabled researchers and practitioners to uncover complex relationships that may remain hidden when using traditional statistical methods. Big data-based workforce analytics frameworks have demonstrated considerable potential for improving organizational forecasting capabilities and supporting proactive human resource management. Research has shown that predictive systems can successfully identify emerging workforce risks and facilitate timely interventions aimed at improving employee retention and organizational effectiveness (Reddy, 2023; Tang et al., 2023; Zheng, 2025).

Machine learning techniques have gained particular attention because of their ability to model nonlinear relationships, capture high-dimensional interactions, and process large and complex datasets. Unlike traditional regression-based approaches, machine learning algorithms are designed to optimize predictive performance rather than merely estimate statistical relationships. Various machine learning methods have been applied to employee turnover prediction, including decision trees, random forests, support vector machines, artificial neural networks, gradient boosting algorithms, and ensemble learning techniques. These approaches have consistently demonstrated superior predictive accuracy compared with many conventional statistical models. As a result, machine learning has become an increasingly valuable tool for organizations seeking to identify turnover risk factors and improve workforce planning processes (Kumar et al., 2023; Lalitha et al., 2023; Márquez et al., 2023).

Recent literature highlights the growing effectiveness of machine learning applications in workforce analytics and employee attrition prediction. Studies conducted across various organizational contexts have demonstrated that predictive models can accurately classify employees according to turnover risk levels and provide valuable insights into the factors contributing to attrition. These studies suggest that variables such as job satisfaction, employee engagement, organizational commitment, compensation, promotion opportunities, workload, work-life balance, managerial support, and performance evaluations often emerge as important predictors of turnover behavior. Furthermore, advances in computational power and algorithmic optimization have enabled the deployment of increasingly sophisticated models capable of achieving high levels of predictive performance. Such developments have encouraged organizations to integrate predictive analytics into broader talent management and strategic workforce planning initiatives (Gazi et al., 2024; Gurung et al., 2024; Nandal et al., 2024).

Among the most promising machine learning techniques for predictive analytics are gradient boosting algorithms, particularly Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and Categorical Boosting (CatBoost). These algorithms belong to the family of ensemble learning methods that combine multiple weak learners to produce highly accurate predictive models. XGBoost has become widely recognized for its computational efficiency, scalability, regularization mechanisms, and predictive accuracy across diverse application domains. LightGBM offers additional advantages through its leaf-wise tree growth strategy and efficient handling of large-scale datasets. CatBoost, meanwhile, has demonstrated exceptional performance when dealing with categorical variables, which are frequently encountered in human resource datasets. Comparative evaluations suggest that these algorithms often outperform traditional machine learning approaches in classification tasks, making them particularly attractive for employee turnover prediction applications (Alqahtani et al., 2024; Chandgadkar et al., 2024; Preena, 2024).

At the same time, Deep Neural Networks (DNNs) have emerged as powerful predictive tools capable of capturing intricate nonlinear relationships within complex datasets. Deep learning architectures have achieved remarkable success across numerous domains, including image recognition, natural language processing, healthcare analytics, and financial forecasting. Within the context of human resource management, neural networks have shown promise for modeling complex employee behaviors and identifying latent patterns associated with workforce outcomes. Some studies have reported strong predictive performance using neural network models for employee turnover prediction, particularly when large datasets are available. Nevertheless, comparisons between deep learning models and advanced gradient boosting algorithms have produced mixed findings, suggesting that the relative effectiveness of these approaches may depend on data characteristics, feature composition, and organizational context. Therefore, additional comparative research is

IJIMOB E-ISSN: 3041-8992

needed to determine the most suitable predictive methods for turnover forecasting in contemporary organizations (Márquez et al., 2023; Qin et al., 2026; Swetha et al., 2022).

Despite substantial progress in predictive modeling, an important challenge remains regarding model interpretability. Many advanced machine learning algorithms are often criticized as "black box" systems because their internal decision-making processes are difficult to understand and explain. In organizational settings, predictive accuracy alone may not be sufficient to support managerial decision-making. Human resource professionals require transparent and interpretable models that explain why certain employees are classified as high-risk turnover cases. Explainable Artificial Intelligence (XAI) has emerged as an important solution to this challenge by providing techniques that reveal the factors influencing machine learning predictions. Methods such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) enable researchers and practitioners to understand the relative importance of predictor variables and interpret individual-level predictions. The integration of explainability with predictive analytics enhances trust, accountability, and practical usability within organizational environments (Akasheh et al., 2024; Alqahtani et al., 2024; Xu, 2026).

The growing interest in explainable machine learning reflects broader trends in organizational digital transformation and responsible artificial intelligence adoption. As organizations increasingly rely on algorithmic decision support systems, concerns regarding transparency, fairness, bias, and ethical decision-making have become more prominent. Explainable models can help organizations identify potential biases, ensure compliance with organizational policies, and provide defensible justifications for human resource interventions. Furthermore, explainable machine learning facilitates communication between technical analysts and organizational stakeholders by translating complex computational outputs into meaningful managerial insights. This capability is particularly important in employee turnover prediction, where intervention decisions may influence employee experiences, organizational culture, and workforce dynamics (Akasheh et al., 2024; Qin et al., 2026; Xu, 2026).

Although previous studies have demonstrated the value of machine learning for turnover prediction, several gaps remain in the literature. First, many investigations focus on a single algorithm rather than conducting comprehensive comparisons across multiple advanced predictive models. Second, some studies emphasize predictive performance while neglecting model interpretability and practical applicability. Third, there remains limited evidence regarding the relative strengths and weaknesses of XGBoost, LightGBM, CatBoost, and Deep Neural Networks within a unified analytical framework. Finally, the rapidly evolving nature of machine learning technologies necessitates ongoing evaluation of emerging methodologies to ensure that organizations adopt the most effective tools for workforce analytics and talent retention. Recent systematic reviews have specifically highlighted the need for comparative research examining state-of-the-art machine learning approaches and their implications for human resource management practice (Alqahtani et al., 2024; Hadiyanto & Anggoro, 2025; Talebi et al., 2025).

Given the strategic importance of employee retention, the increasing availability of workforce data, and the growing adoption of artificial intelligence in human resource management, comparative investigations of explainable machine learning models represent a timely and valuable research direction. Understanding which algorithms provide the highest predictive accuracy while maintaining adequate interpretability can help organizations improve retention strategies, allocate resources more effectively, and enhance workforce stability. Moreover, identifying the most influential turnover predictors can contribute to the development of targeted interventions that address the underlying causes of employee attrition and support long-term organizational success (Gazi et al., 2024; Kundu et al., 2024; Tripathi et al., 2024).

Therefore, the aim of this study was to develop and compare explainable employee turnover prediction models using XGBoost, LightGBM, CatBoost, and Deep Neural Networks and to identify the most influential factors associated with turnover risk through advanced explainable artificial intelligence techniques.

2 Methods and Materials

This study employed a quantitative predictive analytics design using supervised machine learning techniques to develop and compare employee turnover prediction models. The primary objective was to evaluate the predictive performance and interpretability of four advanced machine learning approaches, including Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), Categorical Boosting (CatBoost), and Deep Neural Networks (DNN), for identifying employees at risk of

IJIMOB
E-ISSN: 3041-8992

voluntary turnover. The study was conducted among employees working in medium-sized and large organizations across multiple industries in Canada, including information technology, financial services, healthcare, manufacturing, retail, telecommunications, and professional services. A cross-sectional organizational dataset was assembled through collaboration with human resource departments and organizational data management units.

The final dataset consisted of 4,862 employees from 37 organizations located across the provinces of Ontario, British Columbia, Alberta, Quebec, and Nova Scotia. Employees were included if they had completed at least one year of continuous employment within their respective organizations and had complete personnel records available for analysis. The sample included both employees who remained with their organizations and those who voluntarily left during the observation period. The turnover status of employees served as the target variable and was coded as a binary outcome, where employees who voluntarily resigned were classified as turnover cases and those who remained employed were classified as retention cases. The study incorporated demographic, occupational, performance-related, behavioral, and organizational variables to ensure a comprehensive representation of factors potentially associated with turnover decisions. Prior to model development, data quality assessment procedures were conducted to identify missing values, duplicate records, and outliers. Missing data were handled using multiple imputation and appropriate preprocessing techniques to maximize data integrity and model reliability.

Data were collected using an integrated Human Resource Analytics Framework that combined organizational records, employee survey responses, performance evaluation data, and workforce management indicators. The framework was designed to capture a multidimensional profile of employee characteristics, work experiences, and organizational conditions associated with turnover behavior. Demographic variables included age, gender, educational attainment, marital status, tenure, organizational level, employment status, and department affiliation. Occupational variables consisted of job role, work schedule, promotion history, compensation level, salary growth, training participation, and managerial responsibilities.

Employee engagement was assessed using the Utrecht Work Engagement Scale developed by Schaufeli and colleagues. The instrument measures the dimensions of vigor, dedication, and absorption and has been extensively validated across organizational settings. Responses were recorded on a Likert-type scale, with higher scores indicating greater engagement. Previous studies have consistently demonstrated strong psychometric properties, including satisfactory validity and reliability across diverse occupational populations.

Job satisfaction was measured using the Job Satisfaction Survey developed by Spector. This instrument evaluates multiple aspects of employee satisfaction, including pay, supervision, promotion opportunities, fringe benefits, operating conditions, coworker relationships, nature of work, and communication. Participants rated their level of agreement with survey items using a standardized response scale. The instrument has demonstrated high internal consistency and construct validity in numerous organizational studies.

Organizational commitment was assessed using the Three-Component Organizational Commitment Scale developed by Meyer and Allen. The scale measures affective commitment, continuance commitment, and normative commitment. Higher scores reflect stronger psychological attachment to the organization and a greater likelihood of remaining employed. Previous research has reported robust reliability coefficients and strong evidence of convergent and discriminant validity.

Perceived organizational support was measured using the Survey of Perceived Organizational Support developed by Eisenberger and colleagues. The instrument evaluates employees' perceptions regarding the extent to which the organization values their contributions and cares about their well-being. Higher scores indicate stronger perceptions of organizational support. The measure has been widely utilized in human resource and organizational behavior research and has demonstrated strong psychometric characteristics.

Performance-related variables were extracted from annual performance appraisal records maintained by participating organizations. These indicators included overall performance ratings, goal achievement percentages, competency assessments, productivity metrics, absenteeism records, and disciplinary incidents. Additional workforce analytics variables included overtime hours, training completion rates, internal mobility history, promotion frequency, managerial span of control, employee recognition records, and participation in organizational development initiatives. These objective indicators were integrated with survey-based measures to create a

IJIMOB
E-ISSN: 3041-8992

comprehensive predictive dataset suitable for advanced machine learning analysis.

Prior to model training, all collected variables underwent extensive preprocessing procedures. Continuous variables were standardized using z-score normalization, while categorical variables were encoded using techniques appropriate for each machine learning algorithm. Feature engineering procedures were employed to derive additional predictors, including tenure-to-promotion ratios, engagement-performance interaction terms, absenteeism trends, compensation growth trajectories, and workload balance indicators. These derived variables were included to enhance model performance and capture complex relationships among predictors.

Data analysis was performed using Python-based machine learning libraries, including Scikit-learn, XGBoost, LightGBM, CatBoost, TensorFlow, and Keras. The analytical process involved multiple stages, beginning with exploratory data analysis and feature selection. Descriptive statistics and correlation analyses were initially conducted to examine the distributional characteristics of variables and identify potential multicollinearity issues. Feature importance screening techniques and recursive feature elimination procedures were subsequently applied to identify the most relevant predictors of employee turnover.

The dataset was randomly divided into training, validation, and testing subsets using a 70%, 15%, and 15% split ratio, respectively. Stratified sampling procedures were employed to preserve the proportional representation of turnover and retention cases across all subsets. To address class imbalance, Synthetic Minority Oversampling Technique (SMOTE) and class-weight adjustment methods were implemented where necessary. Hyperparameter optimization was conducted using grid search and Bayesian optimization approaches to identify optimal model configurations for each algorithm.

Four predictive models were developed and compared. The XGBoost model was implemented using gradient boosting decision trees with optimized learning rates, tree depths, and regularization parameters. The LightGBM model utilized histogram-based gradient boosting and leaf-wise tree growth strategies to maximize computational efficiency and predictive performance. The CatBoost model leveraged advanced categorical feature processing techniques and ordered boosting procedures to reduce prediction bias and improve accuracy. The Deep Neural Network architecture consisted of multiple fully connected hidden layers with rectified linear unit activation functions, dropout regularization, batch normalization layers, and adaptive optimization algorithms. Early stopping procedures were implemented to prevent overfitting and improve generalizability.

Model performance was evaluated using multiple classification metrics, including accuracy, precision, recall, F1-score, area under the receiver operating characteristic curve (AUC-ROC), area under the precision-recall curve (AUC-PR), Matthews correlation coefficient, and balanced accuracy. Cross-validation procedures using ten-fold stratified validation were employed to assess model stability and robustness. Comparative analyses were conducted to determine statistically significant differences among model performances.

To enhance interpretability and support practical organizational decision-making, explainable artificial intelligence techniques were incorporated into the analytical framework. SHapley Additive exPlanations (SHAP) were used to quantify the contribution of individual predictors to model outputs and to identify the most influential factors associated with turnover risk. Local Interpretable Model-Agnostic Explanations (LIME) were employed to generate employee-level explanations for specific predictions. Partial dependence plots and accumulated local effect analyses were further utilized to visualize nonlinear relationships between key predictors and turnover probability. These explainability methods enabled both global and local interpretation of model behavior, facilitating a deeper understanding of the mechanisms underlying employee turnover and providing actionable insights for human resource management and organizational retention strategies.

3 Findings and Results

The final dataset consisted of 4,862 employees from 37 Canadian organizations representing multiple industrial sectors. Of the participants, 2,713 (55.8%) were male and 2,149 (44.2%) were female. The mean age of employees was 38.74 years (SD = 9.21), with ages ranging from 21 to 64 years. The average organizational tenure was 6.83 years (SD = 4.57). Most participants held full-time positions (89.6%), while 10.4% were employed on part-time or contract arrangements. Employees from the information technology sector constituted 24.1% of the sample, followed by healthcare (18.7%), financial services (17.5%), manufacturing (15.9%), retail (12.3%), telecommunications (6.2%), and professional services (5.3%). During the

IJIMOB
E-ISSN: 3041-8992

observation period, 917 employees (18.9%) voluntarily left their organizations, whereas 3,945 employees (81.1%) remained employed. Preliminary analyses indicated acceptable distributions across all variables, and no substantial violations of modeling assumptions were identified after data preprocessing, normalization, and outlier treatment procedures.

Table 1

Descriptive Statistics of Major Study Variables

Variable	Mean	SD	Minimum	Maximum
Employee Engagement	3.67	0.81	1.21	5.00
Job Satisfaction	3.54	0.76	1.18	5.00
Organizational Commitment	3.49	0.73	1.14	5.00
Perceived Organizational Support	3.58	0.79	1.09	5.00
Performance Rating	4.01	0.68	1.75	5.00
Absenteeism (Days/Year)	7.92	4.37	0.00	29.00
Training Participation	4.84	2.96	0.00	15.00
Promotion Frequency	1.41	1.07	0.00	6.00
Overtime Hours/Month	10.27	8.51	0.00	54.00
Turnover Probability	0.19	0.39	0.00	1.00

Table 1 presents the descriptive statistics of the principal predictors included in the machine learning models. Employees generally reported moderate to high levels of engagement, job satisfaction, organizational commitment, and perceived organizational support, with mean scores exceeding the midpoint of the scales. Performance ratings were relatively high across organizations, indicating that most employees were evaluated favorably by their supervisors. Absenteeism levels remained relatively low overall, although considerable variability was observed among participants. Training participation and promotion frequency demonstrated substantial dispersion, suggesting notable differences in career development opportunities across organizations. The observed turnover rate of approximately 19% provided sufficient variability for predictive modeling and reflected a realistic organizational retention challenge. The descriptive findings suggest that the dataset captured meaningful variation across both behavioral and organizational characteristics, thereby providing a robust foundation for subsequent machine learning analyses.

Table 2

Comparative Performance of Machine Learning Models

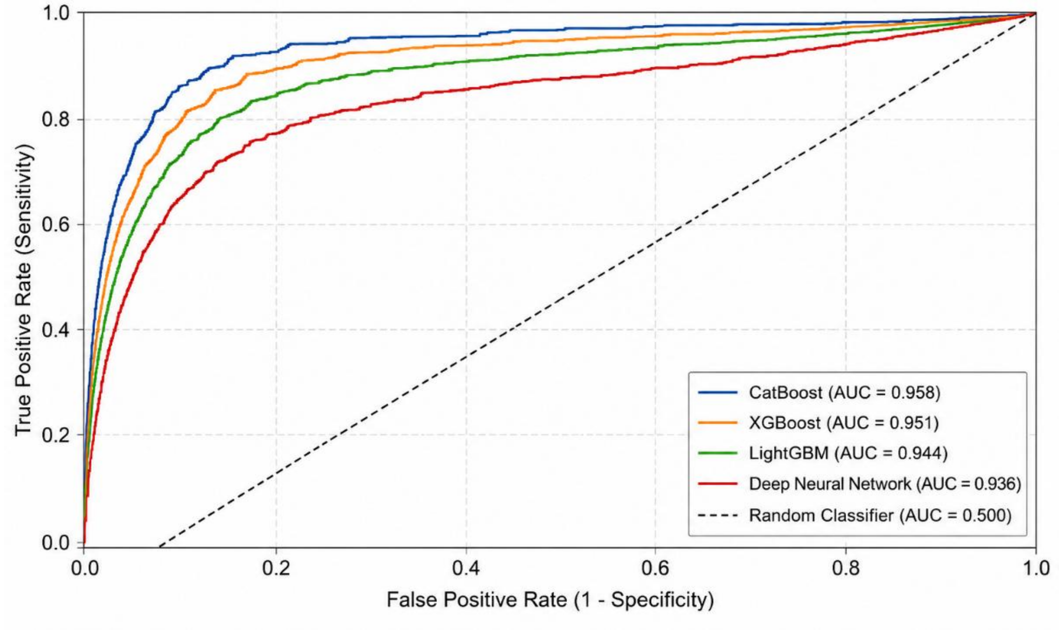
Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC	AUC-PR
XGBoost	0.913	0.887	0.851	0.869	0.951	0.924
LightGBM	0.906	0.878	0.843	0.860	0.944	0.916
CatBoost	0.919	0.895	0.862	0.878	0.958	0.931
Deep Neural Network	0.897	0.864	0.832	0.848	0.936	0.904

The comparative analysis revealed that all machine learning models achieved strong predictive performance; however, important differences emerged among algorithms. CatBoost demonstrated the highest overall performance, achieving an accuracy of 91.9%, an F1-score of 0.878, and an AUC-ROC value of 0.958. XGBoost closely followed, with an accuracy of 91.3% and an AUC-ROC of 0.951. LightGBM also exhibited excellent predictive capability, although its performance was marginally lower than CatBoost and XGBoost across most evaluation metrics. The Deep Neural Network achieved respectable results but consistently underperformed relative to the gradient boosting algorithms. These findings suggest that ensemble boosting methods were particularly effective in capturing complex nonlinear relationships within organizational data. The superior performance of CatBoost may be attributed to its advanced handling of categorical variables and its ability to minimize prediction bias through ordered boosting procedures. Overall, the results indicate that explainable boosting approaches provide highly effective solutions for employee turnover prediction in organizational settings.

IJIMOB
E-ISSN: 3041-8992

Figure 1

Receiver Operating Characteristic (ROC) Curves Comparing XGBoost, LightGBM, CatBoost, and Deep Neural Network Models



The ROC curve analysis further confirmed the superiority of the boosting-based models. Across nearly all classification thresholds, CatBoost maintained the highest true positive rates while simultaneously minimizing false positive rates. XGBoost closely tracked CatBoost's performance, whereas LightGBM demonstrated slightly lower discriminative capability. The Deep Neural Network exhibited the weakest curve among the evaluated models, although its classification performance remained substantially above random prediction. The substantial separation between the ROC curves and the diagonal reference line indicates excellent discriminative power for all models. The observed AUC values exceeding 0.93 across all algorithms suggest that the predictive systems successfully distinguished employees who were likely to leave from those who were likely to remain. These findings demonstrate that machine learning approaches can identify turnover risk with a high degree of accuracy and reliability, making them valuable tools for strategic workforce planning and talent retention initiatives.

Table 3

Top Ten Predictors of Employee Turnover Based on SHAP Analysis

Predictor	Mean SHAP Importance
Job Satisfaction	0.287
Employee Engagement	0.251
Organizational Commitment	0.229
Perceived Organizational Support	0.198
Promotion Frequency	0.175
Salary Growth Rate	0.169
Performance Rating	0.153
Overtime Hours	0.147
Absenteeism	0.138
Training Participation	0.131

The SHAP analysis provided important insights into the drivers of employee turnover and enhanced the interpretability of the predictive models. Job satisfaction emerged as the most influential predictor, followed closely by employee engagement and organizational commitment. Employees reporting low satisfaction, reduced engagement,

IJIMOB
E-ISSN: 3041-8992

and weaker commitment levels exhibited substantially higher probabilities of turnover. Perceived organizational support also played a major role, indicating that employees who felt undervalued or insufficiently supported by their organizations were more likely to resign. Career development indicators, including promotion frequency and salary growth, ranked among the strongest predictors, highlighting the importance of advancement opportunities in employee retention. Interestingly, behavioral indicators such as overtime hours and absenteeism also demonstrated significant predictive value. Employees experiencing excessive workloads or displaying increased absenteeism were more likely to leave their organizations. These findings underscore the multidimensional nature of turnover decisions and demonstrate that both psychological and organizational factors contribute meaningfully to employee retention outcomes.

Table 4

Confusion Matrix Statistics for the Best-Performing CatBoost Model

Classification Outcome	Frequency
True Positives	791
True Negatives	3,842
False Positives	103
False Negatives	126

The confusion matrix analysis of the best-performing CatBoost model revealed highly satisfactory classification accuracy. The model correctly identified 791 employees who subsequently left their organizations and accurately classified 3,842 employees who remained employed. The number of false positive predictions was relatively small, indicating that few employees were incorrectly identified as turnover risks. Similarly, the number of false negatives remained limited, suggesting that the model successfully detected most employees who eventually resigned. From a practical human resource management perspective, these results are particularly valuable because they indicate that organizations can identify a substantial proportion of high-risk employees before turnover occurs. Such predictive capability enables targeted retention interventions, personalized engagement strategies, leadership support initiatives, and career development programs aimed at reducing avoidable employee departures. The strong performance observed in the confusion matrix further reinforces the practical utility and organizational relevance of explainable machine learning models for workforce analytics and talent management.

4 Discussion

The present study sought to develop and compare explainable machine learning models for employee turnover prediction using XGBoost, LightGBM, CatBoost, and Deep Neural Networks and to identify the most influential predictors associated with turnover risk. The findings demonstrated that all four machine learning approaches achieved strong predictive performance, with CatBoost emerging as the best-performing model, followed by XGBoost, LightGBM, and the Deep Neural Network. Furthermore, explainable artificial intelligence analyses revealed that job satisfaction, employee engagement, organizational commitment, perceived organizational support, promotion frequency, salary growth, performance ratings, overtime workload, absenteeism, and training participation constituted the most influential predictors of employee turnover. These findings provide important theoretical and practical insights into the use of advanced analytics for workforce management and contribute to the growing literature on explainable human resource analytics.

One of the most important findings of this study was the superior predictive performance of CatBoost relative to the other machine learning models. The CatBoost algorithm achieved the highest accuracy, precision, recall, F1-score, and AUC-ROC values, indicating a remarkable ability to distinguish employees who were likely to leave from those who were likely to remain within the organization. This finding is consistent with recent studies emphasizing the effectiveness of gradient boosting algorithms for workforce analytics and employee attrition prediction. Research has shown that boosting methods outperform many traditional machine learning techniques because they effectively capture nonlinear interactions among variables and reduce prediction errors through iterative learning processes (Alqahtani et al., 2024; Qin et al., 2026; Talebi et al., 2025). The strong performance of CatBoost may be particularly attributable to its advanced capability for processing

IJIMOB E-ISSN: 3041-8992

categorical variables, which are prevalent in human resource datasets where organizational roles, departments, education levels, and employment categories often represent important predictors. Previous investigations have similarly reported that CatBoost can provide superior predictive accuracy in employee turnover modeling due to its robust handling of categorical information and reduced susceptibility to overfitting (Akasheh et al., 2024; Chandgadkar et al., 2024; Preena, 2024).

The results also demonstrated that XGBoost and LightGBM performed exceptionally well, although slightly below CatBoost. These findings support growing evidence suggesting that gradient boosting methods represent some of the most effective approaches currently available for organizational prediction tasks. Both XGBoost and LightGBM achieved high classification accuracy and demonstrated excellent discriminative capabilities, as reflected in their AUC values. These results align with prior studies that identified boosting algorithms as highly effective tools for employee attrition prediction, workforce planning, and talent retention analysis (Gazi et al., 2024; Kumar et al., 2023; Nandal et al., 2024). The success of these algorithms may be explained by their ability to model complex organizational dynamics that cannot be adequately captured through traditional linear approaches. Employee turnover is inherently multifactorial and often influenced by interactions among psychological, organizational, behavioral, and demographic variables. Boosting algorithms are particularly well suited for detecting these complex relationships, thereby enhancing predictive performance and supporting more accurate workforce forecasting (Reddy, 2023; Tang et al., 2023; Zheng, 2025).

Although the Deep Neural Network demonstrated strong predictive performance, it produced slightly lower classification metrics than the gradient boosting algorithms. This finding contributes to an ongoing debate within the machine learning literature regarding the comparative advantages of deep learning and ensemble-based approaches. While neural networks possess considerable capacity to model nonlinear patterns, they often require larger datasets, extensive hyperparameter optimization, and substantial computational resources to achieve optimal performance. In contrast, boosting algorithms frequently achieve superior results in structured tabular datasets, which characterize most human resource management applications. Similar conclusions have been reported in previous studies comparing deep learning models with advanced ensemble methods in organizational prediction contexts (Márquez et al., 2023; Qin et al., 2026; Swetha et al., 2022). Therefore, the present findings suggest that organizations seeking practical and highly accurate turnover prediction solutions may benefit more from advanced boosting approaches than from purely neural-network-based architectures when working with structured workforce data.

A particularly significant contribution of the present study concerns the identification of the most influential turnover predictors through explainable artificial intelligence techniques. The SHAP analysis revealed that job satisfaction represented the strongest predictor of employee turnover. Employees exhibiting lower satisfaction levels demonstrated substantially higher probabilities of leaving their organizations. This finding is consistent with longstanding turnover theories and empirical research indicating that job satisfaction serves as one of the most reliable antecedents of voluntary turnover. When employees perceive their work experiences negatively, they are more likely to develop withdrawal intentions, seek alternative employment opportunities, and eventually resign. Previous studies examining workforce retention and employee attrition have repeatedly emphasized the central role of job satisfaction in shaping turnover decisions (Dhamija, 2023; Kiran et al., 2023; Tripathi et al., 2024). The current findings reinforce the view that organizations seeking to reduce turnover must prioritize initiatives that enhance employees' overall work experiences and satisfaction levels.

Employee engagement emerged as the second most influential predictor of turnover risk. Employees characterized by lower levels of vigor, dedication, and absorption were more likely to leave their organizations. This result aligns closely with contemporary human resource analytics research demonstrating that engagement functions as a critical determinant of employee retention and organizational commitment. Engaged employees typically experience stronger psychological connections to their organizations, exhibit greater motivation, and demonstrate higher levels of discretionary effort. Consequently, lower engagement often serves as an early warning indicator of potential turnover behavior. Previous workforce analytics studies have reported similar findings and emphasized the importance of engagement monitoring as part of comprehensive retention strategies (Kundu et al., 2024; Ravesangar & Narayanan, 2024; Tripathi et al., 2024). The current results therefore support the integration of engagement metrics into predictive retention systems and organizational decision-making processes.

IJIMOB
E-ISSN: 3041-8992

Organizational commitment and perceived organizational support also emerged as highly influential predictors. Employees who reported weaker emotional attachment to their organizations or perceived lower levels of organizational support were more likely to resign. These findings are theoretically consistent with social exchange perspectives suggesting that employees develop reciprocal relationships with their organizations based on perceived fairness, recognition, and support. When employees believe that their contributions are valued and that the organization genuinely cares about their well-being, they are more likely to maintain long-term employment relationships. Conversely, insufficient support may contribute to psychological withdrawal and turnover intentions. Similar conclusions have been reported in studies emphasizing the importance of supportive organizational environments, effective leadership, and employee-centered management practices for workforce retention (Gazi et al., 2024; Patnaik, 2024; Ravesangar & Narayanan, 2024).

The importance of promotion frequency and salary growth further highlights the role of career development opportunities in turnover decisions. Employees who experienced limited advancement opportunities or slower compensation growth demonstrated significantly higher turnover risk. These findings support theories suggesting that employees evaluate their employment relationships partly through expectations regarding career progression and future rewards. When organizations fail to provide meaningful opportunities for professional growth, employees may seek alternative employment that offers greater developmental prospects. Prior research on HR analytics and retention management similarly identified promotion opportunities and compensation trajectories as critical determinants of employee retention and workforce stability (Kiran et al., 2023; Kumar et al., 2023; Patnaik, 2024). The present findings therefore underscore the importance of strategic talent development programs and transparent career pathways in reducing employee attrition.

Another noteworthy finding concerns the predictive significance of workload-related and behavioral variables, particularly overtime hours and absenteeism. Employees experiencing excessive overtime demands exhibited greater turnover probabilities, suggesting that workload imbalance may contribute to employee dissatisfaction, burnout, and eventual resignation. Similarly, higher absenteeism levels emerged as important indicators of turnover risk, potentially reflecting declining organizational commitment, stress, or disengagement. These findings align with previous studies emphasizing that behavioral indicators often serve as early manifestations of withdrawal processes that precede actual turnover events (Nandal et al., 2024; Reddy, 2023; Tang et al., 2023). The inclusion of these variables within predictive models demonstrates the value of integrating both psychological and behavioral indicators when developing comprehensive turnover forecasting systems.

5 Conclusion

The explainability component of the present study also represents an important contribution to the emerging field of explainable human resource analytics. While many previous turnover prediction studies have focused primarily on predictive accuracy, the current investigation incorporated SHAP-based interpretability techniques that enabled detailed examination of the factors driving model predictions. This approach addresses growing concerns regarding transparency and accountability in algorithmic decision-making. Organizations are increasingly expected to justify data-driven decisions and ensure that predictive systems operate in a fair and understandable manner. Explainable artificial intelligence provides a practical mechanism for achieving these objectives while maintaining high levels of predictive performance. Recent literature has similarly emphasized the strategic importance of explainable analytics for enhancing stakeholder trust, supporting organizational decision-making, and facilitating responsible AI adoption within human resource management contexts (Akasheh et al., 2024; Qin et al., 2026; Xu, 2026).

Several limitations should be considered when interpreting the findings of this study. First, the data were collected from organizations operating within a single national context, which may limit the generalizability of the findings to other countries, cultures, and labor markets. Second, although the dataset was relatively large and diverse, it was based primarily on organizational records and survey measures collected at a specific point in time. Consequently, causal relationships cannot be definitively established. Third, certain potentially relevant predictors, such as personality traits, labor market conditions, leadership styles, and external economic factors, were not included in the analytical framework. Finally, although explainable artificial intelligence techniques improved model transparency, some complexity inherent in advanced machine learning algorithms remains difficult to fully communicate to non-technical stakeholders.

IJIMOB
E-ISSN: 3041-8992

Future studies should examine the performance of explainable machine learning models across different countries, industries, and organizational contexts to evaluate the robustness and generalizability of predictive findings. Longitudinal research designs would be particularly valuable for understanding how turnover risk evolves over time and for identifying dynamic predictors of employee departure. Researchers may also explore the integration of additional data sources, including communication networks, organizational culture indicators, leadership assessments, employee well-being measures, and labor market analytics. Comparative investigations involving emerging artificial intelligence methods, transformer-based architectures, graph neural networks, and hybrid ensemble frameworks may further advance predictive capabilities. Finally, future research should continue examining fairness, bias mitigation, and ethical considerations associated with the implementation of predictive workforce analytics systems.

Organizations should consider adopting explainable machine learning systems as part of their strategic human resource management infrastructure to facilitate proactive identification of employees at risk of turnover. Particular attention should be directed toward monitoring job satisfaction, employee engagement, organizational commitment, and perceived organizational support because these variables emerged as the strongest predictors of turnover risk. Human resource professionals should establish targeted retention programs focused on career development, promotion opportunities, compensation growth, and employee recognition. Organizations may also benefit from routinely monitoring behavioral indicators such as absenteeism and excessive overtime to identify emerging retention challenges before employee departures occur. Importantly, predictive systems should be implemented alongside transparent communication practices and ethical governance frameworks to ensure that data-driven insights are used to support employees rather than merely monitor them. Through the responsible application of explainable analytics, organizations can enhance workforce stability, improve employee experiences, and strengthen long-term organizational performance.

Authors' Contributions

All authors have contributed significantly to the research process and the development of the manuscript.

Declaration

In order to correct and improve the academic writing of our paper, we have used the language model ChatGPT.

Transparency Statement

Data are available for research purposes upon reasonable request to the corresponding author.

Acknowledgments

We would like to express our gratitude to all individuals helped us to do the project.

Declaration of Interest

The authors report no conflict of interest.

Funding

According to the authors, this article has no financial support.

Ethical Considerations

In this research, ethical standards including obtaining informed consent, ensuring privacy and confidentiality were observed.

References

Akasheh, M. A., Hujran, O., Malik, E. F., & Zaki, N. (2024). Enhancing the Prediction of Employee Turnover With Knowledge Graphs and Explainable AI. *IEEE Access*, *12*, 77041-77053. https://doi.org/10.1109/access.2024.3404829

Alqahtani, H., Almagrabi, H., & Alharbi, A. (2024). Employee Attrition Prediction Using Machine Learning Models: A Review Paper. *International Journal of Artificial Intelligence & Applications*, *15*(2), 23-49. https://doi.org/10.5121/ijaia.2024.15202

Chandgadkar, V., Pathak, P., Solanki, A., Shrivastava, A., Maktum, T., & Pulgam, N. (2024). ETPSR: Employee Turnover Prediction With Security Recommendations. https://doi.org/10.21203/rs.3.rs-4261760/v1

Dhamija, S. (2023). Analysing Theoretical Models for Predicting Employee Attrition: A Comparative Study in the FMCG Sector. *Journal of Advanced Zoology*, *44*(S-3), 1179-1191. https://doi.org/10.17762/jaz.v44is-3.1295

Gazi, M. S., Nasiruddin, M., Dutta, S., Sikder, R., Huda, C. B., & Islam, M. Z. (2024). Employee Attrition Prediction in the USA: A Machine Learning Approach for HR Analytics and Talent Retention Strategies. *Journal of Business and Management Studies*, *6*(3), 47-59. https://doi.org/10.32996/jbms.2024.6.3.6

Gurung, N., Gazi, M. S., & Islam, M. Z. (2024). Strategic Employee Performance Analysis in the USA: Deploying Machine Learning Algorithms Intelligently. *Journal of Business and Management Studies*, *6*(3), 01-14. https://doi.org/10.32996/jbms.2024.6.3.1

IJIMOB E-ISSN: 3041-8992

Hadiyanto, H., & Anggoro, Y. (2025). A Skill Matrix Chart Case Study in Global Manufacturing. *Systems Research and Behavioral Science*, *42*(6), 1902-1916. https://doi.org/10.1002/sres.3128

Kiran, P. R., Chaubey, A., & Shastri, R. K. (2023). Role of HR Analytics and Attrition on Organisational Performance: A literature Review Leveraging The SCM-Tbfo Framework. *Benchmarking an International Journal*, *31*(9), 3102-3129. https://doi.org/10.1108/bij-06-2023-0412

Kumar, P., Gaikwad, S. B., Ramya, S. T., Tiwari, T., Tiwari, M., & Kumar, B. (2023). Predicting Employee Turnover: A Systematic Machine Learning Approach for Resource Conservation and Workforce Stability. 117. https://doi.org/10.3390/engproc2023059117

Kundu, B., Shyam, R., & Mondal, B. (2024). The Role of Predictive Analytics in Enhancing Employee Retention in India's Start-Up Ecosystem. *Interantional Journal of Scientific Research in Engineering and Management*, *08*(11), 1-6. https://doi.org/10.55041/ijsrem39218

Lalitha, V., Prithiv, R. S., & Lokesh, P. (2023). A Novel Analysis of Employee Attrition Rate by Maneuvering Machine Learning. *Advances in Science and Technology*. https://doi.org/10.4028/p-sw6mmb

Márquez, B. Y., Realyvásquez-Vargas, A., Lopez-Esparza, N., & Ramos, C. E. (2023). Application of Ordinary Least Squares Regression and Neural Networks in Predicting Employee Turnover in the Industry. *Archives of Advanced Engineering Science*, *2*(1), 30-36. https://doi.org/10.47852/bonviewaaes32021326

Nandal, M., Grover, V., Sahu, D., & Dogra, M. (2024). Employee Attrition: Analysis of Data Driven Models. *Eai Endorsed Transactions on Internet of Things*, *10*. https://doi.org/10.4108/eetiot.4762

Patnaik, T. (2024). The Incorporation of HR Metrics in HR Management: A Theoretical Framework With a Specific Focus on Indian Corporates. *Ajbr*, 5409-5417. https://doi.org/10.53555/ajbr.v27i3s.3115

Preena, R. E. (2024). Techniques to Predict Employee Attrition Using Optimized Levy Fruit Fly Optimization Algorithm. https://doi.org/10.21203/rs.3.rs-4127736/v1

Qin, R., Qi, X., Yuan, Y., & Alataş, B. (2026). Employee Turnover Prediction Research of Human Resource Management on Machine Learning Algorithms and Big Data Analysis. *Journal of Organizational and End User Computing*, *38*(1), 1-27. https://doi.org/10.4018/joeuc.399146

Ravesangar, K., & Narayanan, S. (2024). Adoption of HR Analytics to Enhance Employee Retention in the Workplace: A Review. *Hum. Resour. Manag. Serv.*, *6*(3), 3481. https://doi.org/10.18282/hrms.v6i3.3481

Reddy, J. S. (2023). Big Data-Based Framework for Prediction of Employee Attrition by Using Deep Data People Analytics. https://doi.org/10.52783/jier.v3i2.432

Swetha, M., Mahalakshmi, S., Pushpa, S., Madihalli, A. T., Ananya, A., & Bhardwaj, A. (2022). Machine Learning Approach for Forecasting Job Appeasement and Employee Corrosion. https://doi.org/10.15439/2023r17

Talebi, H., Bardsiri, A. K., & Bardsiri, V. K. (2025). Machine Learning Approaches for Predicting Employee Turnover: A Systematic Review. *Engineering Reports*, *7*(8). https://doi.org/10.1002/eng2.70298

Tang, M., Zhao, T., Hu, Z., & Li, Q. (2023). Research on Risk Prediction and Early Warning of Human Resource Management Based on Machine Learning and Ontology Reasoning. *Tehnicki Vjesnik - Technical Gazette*, *30*(6). https://doi.org/10.17559/tv-20230810000869

Tripathi, A. K., Kumari, P., Katiyar, P., Chaudhary, V. S., & Sharma, M. (2024). Workforce Attrition and Engagement Analysis Using Machine Learning. https://doi.org/10.55524/csistw.2024.12.1.26

Xu, Y. (2026). Improving Employee Retention With Machine Learning for Proactive Turnover Risk Detection and Strategic Interventions. https://doi.org/10.21203/rs.3.rs-8378596/v1

Zheng, X. (2025). Construction of a Big Data-Driven Predictive Analysis Platform for Hospital Talent Attrition. *Blockchain in Healthcare Today*, *8*(3). https://doi.org/10.30953/bhty.v8.433

IJIMOB
E-ISSN: 3041-8992